(12) UK Patent Application (19) GB (11) 2 398 452 (13) A

(43) Date of A Publication 18.08.2004

(21) Application No: 0303638.1

(22) Date of Filing: 17.02.2003

(51) INT CL[7]:
H04N 1/42 , H04B 1/74 , H04L 1/22 , H04M 3/22

(52) UK CL (Edition W):
H4K KOD3
H4L LFMA

(56) Documents Cited:
GB 2366703 A

(58) Field of Search:
INT CL[7] H04B 1/74, H04L 1/22, H04M 3/22, H04N 1/42
Other: ONLINE : EPODOC ; WPI ; PAJ

(71) Applicant(s):
NEC Technologies (UK) Limited
(Incorporated in the United Kingdom)
Level 3, Imperium, Imperial Way,
READING, Berks, RG2 0TD,
United Kingdom

(72) Inventor(s):
Simon Fabri

(74) Agent and/or Address for Service:
W P Thompson & Co
55 Drury Lane, LONDON, WC2B 5SQ,
United Kingdom

(54) Abstract Title: **Method and apparatus for controlling video telephony communications**

(57) The invention provides for a method of controlling video telephony communications between communication terminals so as to allow for switching between a video telephony service and a voice-only service, wherein the said step of switching is initiated by monitoring the quality of video output at the at least one of the terminals so as to determine a deterioration in the quality of said video output, and in which the service can return to video telephony when appropriate.

GB 2 398 452 A

Fig. 1

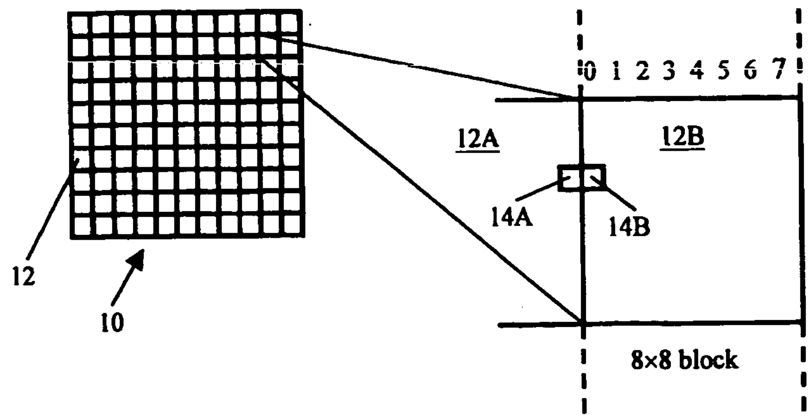


Fig. 2A

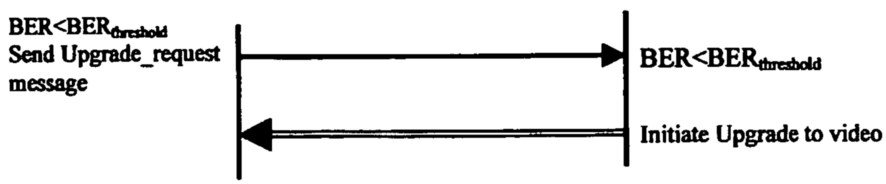
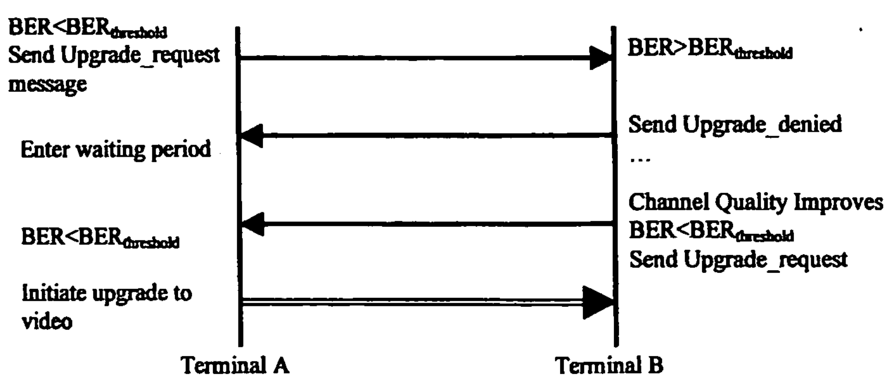
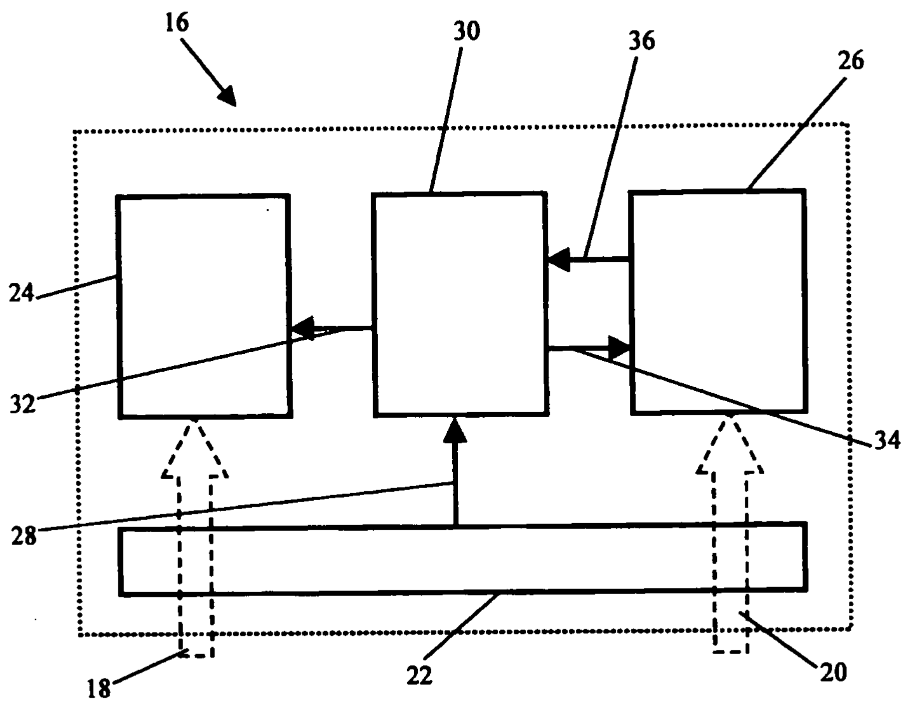


Fig. 2B



2/2

Fig. 3



Method and Apparatus for Controlling Video Telephony Communications

The present invention provides for a method and related apparatus for controlling video telephony communications and in particular, but not exclusively, for controlling video telephony communications over a mobile communications network.

Many mobile networks and terminals currently exhibit the ability to support both speech-only telephony and video telephony. Although video telephony can significantly enhance communications between two participants, it is still accepted that the speech is generally perceptually more important than the corresponding video. It is indeed near impossible to achieve a viable two-way communication if no speech or text path is maintained. Moreover, compressed video signals place significant demands upon the underlying transport networks since the throughput requirements are higher, and compressed video is considerably more susceptible to network errors. Situations therefore often arise in which the available channel conditions are too poor to support video of an acceptable quality, but which conditions are however sufficient for voice telephony. In such situations, it is preferable to switch automatically to speech services, so that at least the conversation may be maintained. When the channel conditions improve, the terminals can revert back to voice telephony.

There currently exist various methods for implementing service changes between voice-only and video telephony communications, and these are largely dependent upon the underlying network topology. For circuit-switched UMTS video telephony, the service change feature may be handled in a proprietary manner solely by the terminal. In this case, no intelligence is required in the network, and either one or both of the terminals in the two-way communication exchange can initiate a switch from a video call to a voice-only call or vice versa. A second approach makes use of network mechanisms for

managing bearer modification procedures. One such method is known as Service Change and UDI/RDI Fallback (SCUDIF) [3GPP TS 23.172] which describes methods by which the network users may carry out a service change between voice services using the standard UMTS/GSM speech bearers and video telephony using UDI/RDI CS bearers.

The present invention seeks to provide for a method and related apparatus for controlling video telephony communications and having advantages over known such methods and apparatus. In particular, the present invention seeks to provide for means for controlling switching between a video telephony service and a voice-only service within a mobile communications device terminal.

According to one aspect of the present invention there is provided a method of controlling video telephony communications between communication terminals so as to allow for switching between a video telephony service and a voice-only service, wherein the said step of switching is initiated responsive to a step of monitoring the quality of video output at at least one of the terminals so as to determine a deterioration in the quality of said video output.

As will therefore be appreciated, the present invention provides for a mechanism for automatically switching between speech-only and video telephony services, and vice versa, over mobile networks. The method is advantageously network-adaptive in that if, while in a video telephony call, the transmission conditions deteriorate to such an extent that the received picture quality is very poor, the terminals will automatically switch to a more robust voice-only service. If the conditions over the network improve, the service can readily revert to video telephony.

Advantageously a deterioration in the quality of the video output comprises determining an average of the amplitude difference between the edge pixels of adjacent blocks within a video frame.

Advantageously the step of determining the average amplitude difference is for a whole frame, and in particular over a plurality of frames.

Alternatively, the method includes the step of identifying a deterioration in the quality of the video output on the basis of the number of corrupted video blocks, or at corrupted portions of a video frame.

According to another aspect of the present invention there is provided a mobile communications terminal including means for switching between a video telephony service and a voice-only service, the terminal further including means for monitoring the quality of video output at the terminal so as to identify deterioration in the quality of said video output and thereby activate the said means for switching.

The invention is described further hereinafter, by way of example only, with reference to the accompanying drawings in which:

Fig. 1 is a schematic representation of a video frame as employed within the present invention;

Figs 2a and 2b illustrate the determination of a possible upgrade from a speech only service to a video-telephony service; and

Fig. 3 is a schematic block diagram of a mobile terminal according to an embodiment of the present invention.

In illustration of an embodiment of the present invention reference is made to a situation in which two mobile communications network users are

participating in a video telephony session using mobile terminals. If either one of the links is experiencing high bit-error rates, there will be a degradation of video quality. Although one criterion that could be used to make a decision to switch from video to speech-only is the received bit error rate, there are however a number of limitations to this technique. Each mobile terminal can only estimate the bit error rate on its own air interface and so has no visibility of the errors that occur on an end-to-end basis during the telephony session. Measuring the error rate on an end-to-end basis would require inserting extra channel coding mechanisms within the audio/video multiplexing arrangement, which could have disadvantageous implications upon interoperability with other terminals. Moreover, the raw bit error rate is only one of the aspects affecting video quality, and the relationship between the bit error rate and the bit error rate and the resulting video quality is very difficult to model. For example, the effect of jitter arising from variations in transit delay, and lost packets which may occur when transiting over packet-based core networks and during handover, also may have significant impact upon the received video quality.

The present invention makes use of a technique designed to assess the quality of received MPEG-2 video. In a paper by Lauterjung of Rohde & Schwarz, a parameter called Digital Quality Level (DVQL-W) is introduced. In this paper it is argued that the most significant impairment introduced by the MPEG-2 compression algorithm is the blocking effect caused by the basing of the encoding process upon "blocks" of 8x8 pixels or "macroblocks" of 16x16 pixels. A technique for measuring this "blockiness" is presented in this paper and is based on calculating the average of the amplitude differences between the edge pixels of adjacent blocks and macroblocks. This average value produces a metric which has been shown to correspond closely to the perceived video quality, and has been used to develop equipment for assessing the quality of the received MPEG-2 transmission.

In mobile transmission scenarios, most visual impairments are caused by transmission errors. In all motion-compensated transform-based predictive codecs, this results in errors in motion prediction, and errors in transform decoding and variable length code decoding. When the errors are detected by the decoder, error concealment algorithms may be employed. However, as all blocks are encoded separately, the channel errors affect each block separately, such that one 8x8 block may be totally corrupted, whereas an adjacent block may remain totally unaffected. The resulting visual effect of distortion may therefore be measured by way of a determination of the amplitude differences between adjacent blocks.

Turning now to Fig. 1, there is illustrated a video frame 10 composed of a plurality of blocks 12, one adjacent pair of which 12A, 12B is illustrated within an enlarged portion of that figure and which illustrates each block as being an 8x8 pixel block.

Adjacent edge pixels 14A and 14B of each of the respective two blocks 12A, 12B are shown and it is the amplitude differences between such adjacent pixels of adjacent blocks that are determined in order to arrive at an average amplitude difference between the adjacent blocks such as 12A and 12B.

It is shown in the paper by Lauterjung that the average amplitude difference of adjacent blocks for an entire frame n can be computed as ADn, and a moving average of these values over N frames can be measured. The value of N will depend upon the frame rate of the received value, and the resulting metric, VideoDistortion_N, will then represent the average distortion of the received video sequence over the past N frames. If this value exceeds a threshold value, this can be taken as an accurate indication that transmission effects have degraded the video quality to such an extent that a service change mechanism should be initiated so as to downgrade the service from video telephony to voice only. The threshold may be preset in the terminal, or may be "learnt" in an adoptive manner. The learning process could for

example involve the user manually downgrading from a video telephony service to a voice-only service a number of times, with the value of $VideoDistortion_N$ being measured at the instant of each downgrade decision. An average value of the metric can then be used as the threshold Video value for initiating subsequent video downgrades.

An alternative embodiment of the present invention uses the number of detected corrupted video blocks (determined by blocks which cannot be correctly decoded) as signalled by the video decoder, averaged over N frames. This method may not however match the user's perception as accurately as the average distortion mentioned above since the video decoder may not classify all corrupted blocks as being in error, as many of the corruptions will result in a syntactically correct bitstream. However this alternative embodiment is disadvantageously simpler to implement and has a lower processing overhead.

In any case, since block-based algorithms are used by all commonly used standards and proprietary codecs, including (MPEG-2, MPEG-4,H.263,WMV), the technique of the present invention can advantageously be adapted for use in any video telephony system. Moreover, since the method of the invention is effectively based on an objective assessment of the received video quality, it is a more reliable technique than any algorithm based on the monitoring of network parameters.

When deciding to upgrade from a voice-only service to a video parameters telephony service, the assessment of received speech quality is not available for use as a reliable metric to decide whether to upgrade back to video telephony service. The difficulty arises when attempting to assess, in a reliable way, the quality of received compressed speech without having access to the original source speech. One of the reasons for this difficulty is the wide range of different compression technologies used by different compression algorithms. For example one algorithm, the Vqmon/EP by Telchemy, which is

used for object speech assessment single-stimulus i.e. with no information about the original speech, serves to map the received traffic characteristics onto an estimated resulting Mean Opinion Score (MOS). This mapping is different for each decoder combination used, and is optimised for Voice over IP applications.

In view of such factors, the illustrated embodiment of the invention is arranged to employ a "Dual Decision" whereby the bit error rate (BER) over the local radio interface is measured by each terminal over a plurality of time frames, N_{speech_frames}, by both terminals and an average BER for the measurement period is computed. An example is now illustrated with reference to Figs. 2A and 2B and which involves two terminals A and B. If the measured bit error rate at terminal A is lower than $BER_{threshold}$, terminal B will initiate the upgrade from the voice-only service to the video telephony service. If however terminal B replies in the negative, or does not reply, the upgrade to the video telephony service will not take place. Terminal A will enter a waiting state for a preset period $Wait_A$, during which it may not request another upgrade until the waiting period is over, although it will continue to carry out bit error rate measurements. If however terminal A receives a request from terminal B, and the current averaged BER measurement at terminal A is still lower than $BER_{threshold}$, then terminal A will initiate the upgrade to the video telephony service.

In Fig. 3 there is illustrated a schematic representation of a mobile radio communications terminal 16 arranged for operation in accordance with the present invention.

The terminal 16 is arranged to receive a speech telephony data stream 18 and a video telephony data stream 20, although only one of the said streams is active within the terminal at any one time.

Both the speech telephony data stream 18 and the video telephony data stream 20 are delivered via a communications protocol stack 22 to a speech telephony application 24 and a video telephony application 26 respectively.

The speech telephony application 24 includes voice encoding, decoding and related input and output of the speech signals, whereas the video telephony application 26 includes voice and video encoding, decoding and related input and output signals.

The communications protocol stack 22 is also arranged to deliver a communications channel quality measurement signal 28 to a switch arrangement 30.

The switch arrangement 30 is arranged to deliver switching control signals 32, 34 to the speech telephony application 24 and the video telephony application 26 respectively.

The video telephony application 26 also includes the functionality for assessing the quality of the video output at the terminal and delivers an appropriate video quality signal 36 to the switch arrangement 30. Such functionality can be based either on the edge pixel amplitude measurements, or the corrupted block measurements discussed above or otherwise.

In accordance with the present invention, should the perceived video quality output at the terminal deteriorate as indicated by the signal 36 from the video telephony application 26, the switch arrangement 30 is arranged to deliver a switching decision signal via the decision signal lines 32,34 to effectively activate only the speech telephony application 24 and its related speech telephony data stream 18.

Should at some future stage the communications protocol stack 22 identify an improvement in bit error rates, the channel quality measurements signal 28 effectively serves to control the switch arrangement 30 to return the video telephony data stream 22 to an active state so as to reintroduce voice and video communication.

Of course, the invention is not restricted to the details of the foregoing embodiment and, in particular, any appropriate form of upgrading to a video telephony service can be employed.

Claims

1. A method of controlling video telephony communications between communication terminals so as to allow for switching between a video telephony service and a voice-only service, wherein the said step of switching is initiated responsive to a step of monitoring the quality of video output at at least one of the terminals so as to determine a deterioration in the quality of said video output.

2. A method as claimed in Claim 1, wherein a deterioration in the quality of the video output at the at least one of the terminals is identified by identifying pixel blocks within a video frame and determining an average of the amplitude difference between edge pixels of adjacent blocks, and wherein a switch to a voice only service is made responsive to the said amplitude difference.

3. A method as claimed in Claim 2, and including the step of determining the average amplitude difference for an entire frame.

4. A method as claimed in Claim 3, and including the step of determining a moving average of the said difference over a plurality of frames.

5. A method as claimed in any one of Claims 2, 3 or 4, and including the step of comparing the value derived from the amplitude difference with a threshold value within a terminal.

6. A method as claimed in Claim 1 wherein a deterioration in the quality of the video output at the at least one of the terminals is

identified on the basis of a determination of the number of detected corrupted video blocks within a frame.

7. A method as claimed in Claim 1, wherein a deterioration in the quality of the video output at the at least one of the terminals is identified on the basis of a determination of corrupted portions of a video frame.

8. A method as claimed in Claim 6 or 7 wherein the corrupted blocks or portions of a video frame are averaged over a plurality of frames.

9. A method as claimed in Claim 6, 7 or 8, wherein the corrupted blocks or portions are identified in a video decoder.

10. A method as claimed in any one or more of Claims 1 to 9 and including the step of controlling a return to video telephony service responsive to bit error rates identified in the voice only signal.

11. A method as claimed in Claim 10 and including the step of determining the bit error rate at each of the said terminals.

12. A method as claimed in Claim 11 and including the step of initiating a return to a video telephony service if the bit error rate value measured at each of the said terminals is above a threshold value.

13. A method as claimed in Claim 10, 11 or 12, wherein the measured bit error rate comprises an average value taken over of plurality of frames.

14. A mobile communications terminal including means for switching between a video telephony service and a voice-only service, the terminal further including means for monitoring the quality of video output at the terminal so as to identify deterioration in the quality of said video output and thereby activate the said means for switching.

15. A terminal as claimed in Claim 14 wherein the said means for determining the deterioration in the quality of the video telephony service is arranged to identify pixel blocks within a video frame, and to determine the average amplitude differences between edge pixels of adjacent blocks so as to control the said means for switching responsive to the said amplitude differences.

16. A terminal as claimed in Claim 15, and including means for determining the average amplitude difference for an entire frame.

17. A terminal as claimed in Claim 16, and including means for determining a moving average of the said difference over a plurality of frames.

18. A terminal as claimed in any one of Claims 15, 16 or 17, and including means for comparing the value derived from the amplitude difference with a threshold value.

19. A terminal as claimed in Claim 14, wherein the said means for monitoring the quality of the video output comprise means for identifying the number of corrupted video blocks within a frame.

20. A terminal as claimed in Claim 14, wherein the said means for monitoring the quality of the video output comprises means for identifying corrupted portions of a video frame.

21. A terminal as claimed in Claim 19 or 20, wherein the corrupted blocks or portions of a video frame are averaged over a plurality of frames.

22. A terminal as claimed in any one or more of Claims 14 to 21, and including means for controlling a return to a video telephony service responsive to bit error rates identified in the voice only signal.

23. A terminal as claimed in Claim 22 including means for identifying the bit error rate at each of the said terminals.

24. A terminal as claimed in Claim 23 including means for initiating a return to a video telephony service if the bit error rate value measured at each of the said terminals is above a threshold value.

25. A terminal as claimed in Claim 22, 23 or 24, and including means for taking an average of the bit error rate over a plurality of frames.

26. A communications network having a plurality of terminals as defined in any one or more of Claims 14 to 25.

27. A communications network having means for controlling video telephony communications in accordance with the steps of any one or more of Claims 1 to 13.

28. A method of controlling video telephony communications between communication terminals substantially as hereinbefore

described with reference to, and as illustrated in, Fig. 1, Figs 2A, 2B and Fig. 3 of the accompanying drawings.

29. A mobile communications terminal substantially as hereinbefore described with reference to, and as illustrated in Fig. 1, Figs. 2A, 2B and Fig. 3 of the accompanying drawings.

30. A communications network substantially as hereinbefore described with reference to, and as illustrated in, Fig. 1, Figs. 2A, 2B and Fig. 3 of the accompanying drawings.

The Patent Office

INVESTOR IN PEOPLE

Application No:	GB 0303638.1	**Examiner:**	Richard Howe
Claims searched:	1-30	**Date of search:**	6 May 2003

Patents Act 1977 : Search Report under Section 17

Documents considered to be relevant:

Category	Relevant to claims	Identity of document and passage or figure of particular relevance
X	1,14,26, 27	GB 2 366 703 A NEC Corporation - see whole document

Categories:

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.

Field of Search:

Search of GB, EP, WO & US patent documents classified in the following areas of the UKC^{V}:

None

Worldwide search of patent documents classified in the following areas of the IPC^{7}:

H04N (1/42) ; H04M (3/22) ; H04B (1/74) ; H04L (1/22)

The following online and other databases have been used in the preparation of this search report :

EPODOC ; WPI ; PAJ